RESEARCH DEPARTMENT

Open - limit calculus

RESEARCH REPORT No. F-026

UDC 517.2/.3 : 517.432.1 1966/54

THE BRITISH BROADCASTING CORPORATION
ENGINEERING DIVISION

RESEARCH DEPARTMENT

OPEN-LIMIT CALCULUS

Research Report No. F-026
UDC 517.2 /3 : 517.432.1 1966/54

J.W. Head, M.A., F.Inst.P., A.M.I.E.E., F.I.M.A.

Head of Research Department

Research Report No. F-026

OPEN-LIMIT CALCULUS

September 1966

Research Report No. F-026
UDC 517.2/.3 : 517.432.1 1966/54

OPEN-LIMIT CALCULUS

SUMMARY

Given a function of time which is "admissible" for the purpose of describing natural phenomena, its derivative and/or integral may not be "admissible". From the point of view of an engineer or scientist, this is of small importance, since neither pure integration nor pure differentiation occurs in natural phenomena. For example, capacitances and inductances always appear in circuits in association with at least residual shunt and series resistance. From the point of view of the pure mathematician, a function "admissible" for describing natural phenomena is also a function having a Fourier transform - the necessary qualifications are the same. The difficulty in pure mathematics, however, is elementary but fundamental and far reaching. It is associated with a wrong application of basic mathematical methods. A process or quantity derived by closing limits we here call "nominal". If two or more such processes or quantities occur simultaneously in a calculation, it is sometimes critically important which limit is closed first. If the limits are closed in the correct order, no difficulty will occur, and with the Fourier integral no special techniques will be required. Due attention must be paid to the possibility that a problem may be so formulated that a critically important limit has been inadvertently closed prematurely.

1. INTRODUCTION

A student usually first meets the idea of a limit in connection with the process of differentiation. The idea of "average speed" has been familiar at a much earlier age, but now the concept of "instantaneous speed" is introduced. This concept is very difficult to get across to an individual, and still more to a class, when they meet it for the first time. The students concerned are quite right to put up a kind of unconscious resistance to the idea, for the concept of a closed limit (in many apparently different forms, such as absolute zero, infinity, "pure" differentiation, integration, resistance, inductance, capacitance etc) is responsible for many unnecessary difficulties in the study of natural phenomena. These difficulties are not inherent in the phenomena, but in the manner of thought associated with these concepts.

We must be clear from the beginning that there are two complementary and fundamentally different ways of specifying what we wish to discuss - a "metrical" way, analogous to the "average" speed mentioned above, and a "nominal" way, corresponding to the idea of "instantaneous" speed. "Metrical" and "nominal" specifications are therefore fully compared and contrasted in Section 2 which follows.

2. "METRICAL" AND "NOMINAL" SPECIFICATION

Examples of "nominal" specification in ordinary life are: a loaf of bread, a dozen eggs, £1, a one-inch board or a one-inch diameter shaft. These nominal specifications can be used only when no critical comparison of items is envisaged; they vary widely in their accuracy and their validation. A loaf of bread is validated by legal penalties for short-weight; a £1 note has no definite buying power but is legal tender; a steel rod having a nominal diameter of 1 inch might have an actual diameter specified ("metrically") as being between 0·999 and 1·001 in. But such "nominal" specifications would be useless on an assembly line. "Metrical" specification is then required, and equations must be replaced by inequalities. Thus if a one-inch shaft has to be fitted into a one-inch collar, we need a "metrical" specification that the diameter d inches of the shaft satifies

$$1 - \eta - \epsilon < d < 1 - \eta \qquad \epsilon, \eta > 0 \qquad (1a)$$

while the diameter D in. of the collar satifies

$$1 < D < 1 + \xi \qquad \xi > 0 \qquad (1b)$$

To secure a fit, the critically important point is that the greatest permitted d be less than the least permitted D. If the sign of equality occurs, it indicates usually that "nominal" specification is being used. Occasionally it indicates that a change of nomenclature is convenient, as in the identity

$$a^2 - b^2 = (a - b)(a + b) \tag{1c}$$

in which both sides are alternative names for the same entity or thing. It may happen that either name is the more convenient; both are exactly the same. We can speak of a nominal specification as "exactly" 12 eggs or "exactly" £1, but a metrical specification of "exactly" 12 inches is nonsense, since this really means 12·00000 . . . . inches and requires infinite information.

Although in science and technology nominal values can thus only be used when there is no critical comparison, it is always possible to unclose a prematurely-closed limit if we are unexpectedly confronted with a critical comparison. The expression py or dy/dt is a nominal or conceptual derivative of y; it can be regarded as derived by the closing of any one of the three limits

$$\lim_{\delta t \to 0} \frac{\delta y}{\delta t}; \quad \lim_{\epsilon \to 0} (\mathrm{p} + \epsilon); \quad \lim_{\nu \to 0} \left(\frac{1}{\mathrm{p}} + \nu\right)^{-1} \tag{2}$$

each of which is a metrical entity before the limit is closed. Again, we use freely the nominal or conceptual ideas of "pure" capacitance, "zero" impedance or admittance, and all is well as long as we always remember the existence of at least residual associated resistance (as in Fig. 1 discussed below).

In pure mathematics, however, a "nominal" specification implies that the limit has been closed. If we are given that

$$b = a \text{ and } c = a \tag{3}$$

the implication is

$$b = c \tag{4}$$

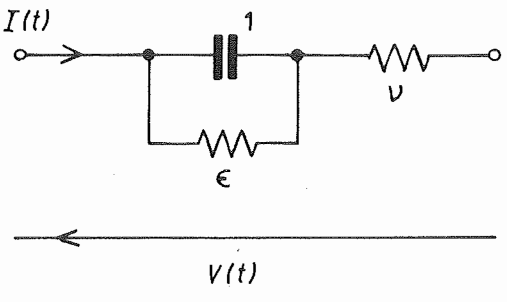

Fig. 1 - Residual elements associated with a unit capacitance

and no information is available about the signs of any of the quantities $(b - c)$, $(c - a)$ and $(a - b)$. This information may be critically important, as in the case of the shaft and collar already discussed.

If ϵ is arbitrarily small and positive, the equation

$$b = a + \epsilon$$

is in effect a shorthand for the inequalities

$$a < b < a + \epsilon$$

and it is possible to retain the exactness of pure mathematics if we refrain from closing limits too soon. We may even have to "unclose" them, as the conceptual expression dy/dt or py was "unclosed" by means of the expressions (2).

3. THE ENGINEER'S POINT OF VIEW

Let us first consider what restrictions must be put upon a function (of time) so that it can be regarded as "admissible" for the purpose of describing natural phenomena. We must then consider how the repeated application of some of the above-mentioned conceptual processes, like pure differentiation, may turn an "admissible" function into an "inadmissible" one. In some situations we can, and frequently do "get away with" the use of closed limits, especially if there is only one conceptual or closed-limit operation involved in the problem under discussion. But when two such entities are involved simultaneously, as in the famous question "What happens when an irresistible force is applied to an immovable body?", we have a situation about which it is impossible to talk sense because the situation is inadequately specified. We must determine "metrical" processes analogous to the above-mentioned conceptual processes, and differing from them to an arbitrarily small extent, such that the application of any combination of these "metrical" processes to an "admissible" function always gives an "admissible" result. Success in this task means that we need never bother about mathematical complications and monstrosities such as functions having an infinite number of discontinuities in a finite range, discontinuities more complicated than an isolated step, Lebesgue integration, distribution theory, etc. Such matters can be relegated to the status of a game, analogous to chess - very enjoyable for those who like mathematical abstractions for their own sake, without caring how little they make the work of others easier.

4. ADMISSIBLE FUNCTIONS OF TIME

A minimum qualification for an "admissible"

function $h(t)$ of time might be expressed in the form that $h(t)$ is everywhere finite, and is of limited extent. This means either that $h(t)$ is zero when t is outside certain limits, or, if $h(t)$ is different from zero for all t, then $h(t)$ must tend to zero as t tends to $\pm\infty$. Furthermore, $h(t)$ must have an ultimately smooth fine structure. However, these points are most easily specified and understood with the help of Fourier theory. In the language of this theory, $h(t)$ must have no components at zero frequency or at infinite frequency.

By definition, if $F(\mathrm{p})$ is the Fourier transform of $h(t)$

$$h(t) = \frac{1}{2\pi \mathrm{j}}\int_{-\mathrm{j}\infty}^{\mathrm{j}\infty} e^{\mathrm{p}t}F(\mathrm{p})\,d\mathrm{p}; \qquad F(\mathrm{p}) = \int_{-\infty}^{\infty} e^{-\mathrm{p}t}h(t)\,dt \tag{5}$$

Both integrals in (5) are infinite integrals, so that their evaluation involves the closing of limits associated with p and t respectively. A necessary condition of convergence and hence usefulness of these integrals is

$$F(\mathrm{p}) \to 0 \text{ as } \mathrm{p} \to \pm \mathrm{j}\infty \tag{6a}$$

$$h(t) \to 0 \text{ as } t \to \pm\infty \tag{6b}$$

Condition (6*a*) ensures that

$$h(t) = \tfrac{1}{2}\lim_{\epsilon\to 0}\,[h(t+\epsilon) + h(t-\epsilon)] \tag{7}$$

while conditions (5) ensure that $F(\mathrm{p})$ remains bounded as p tends to zero. Note that proceeding to the limit for p tending to $\pm\infty$ sometimes presents great difficulty and sometimes ambiguity. Sometimes there appears to be no definite limit. Closer inspection reveals, however, that in such cases a second quantity or process, derived by closing a limit, is involved. The solution lies in "unclosing" this prematurely closed limit before attempting to close the limit for p.

5. APPLICATION OF CONCEPTUAL PROCESSES TO ADMISSIBLE FUNCTIONS

Suppose that a particular function $h(t)$ qualifies for "admissibility" by tending to zero as t tends to $\pm\infty$ and that it has a Fourier transform $F(\mathrm{p})$, defined by (5) such that F(O) has a finite value a while for large values of p, the expression $\mathrm{p}F(\mathrm{p})$ tends to a constant limit K. Then if we differentiate $h(t)$, the Fourier transform of the derivative $h'(t)$ is $\mathrm{p}F(\mathrm{p})$, which tends to zero with p but tends to K, not zero, when p tends to infinity, so that $h'(t)$ fails to qualify as an "admissible" function. Again if $I(t)$ is the integral of $h(t)$ (with lower limit of integration $-\infty$) the Fourier transform of $I(t)$ is $F(\mathrm{p})/\mathrm{p}$. This expression tends to zero as p tends to infinity, but tends to infinity as p tends to zero, so that $I(t)$ is again an "inadmissible" function in spite of the fact that $h(t)$ is "admissible". Clearly admissible functions $h(t)$ which can be repeatedly differentiated and/or integrated are few and peculiar.

Now the way to overcome these difficulties is to replace conceptual differentiation, integration, etc. by analogous "metrical" processes as indicated in Section 6 below in a critical situation requiring accurate and detailed specification and calculation. Many situations, however, are not critical in this sense, and for these situations nominal or quantized concepts greatly simplify the relevant mathematical work. In electrical technology, likewise, the concepts of pure resistance, capacitance, etc. are often extremely helpful as long as they are used with care and we never allow ourselves to forget their idealized nature.

6. METRICAL DIFFERENTIATION AND INTEGRATION

Now the cure for the difficulties associated with conceptual differentiation, integration etc. is best found by means of physical intuition. If we replace the conceptual idea of "instantaneous velocity" by that of "average velocity over the shortest time-interval that we can measure in the situation under discussion", we know exactly what we are talking about. Likewise, a nominally "pure" unit capacitance in practice necessarily has at least residual series resistance (say ν ohms) and shunt conductance (say ϵ mhos) associated with it. It is thus better represented in terms of Figure 1.

In science (the study of observables), all actual observations are averages, probabilities and similar quantities. Instantaneous velocity, position etc. are not observed or observable. We expect that the simplest formulation of our observations will be in terms of ideal concepts like Maxwell's equations. But using such idealizations involves "contraction", of information, by the formation of various mean or average values, for example, contraction of tensors, mean values of vectors, scalar products of vectors, formation of moments, centroids, frequency components, etc.

Ideally, a pure unit capacitance can be regarded as an integrator of all current $I(t)$ applied to it, and as a differentiator of all voltage $V(t)$ applied to it; both ways of regarding it are equivalent expressions of what a capacitance does. When the residual elements are taken into account, however, the essential fact is that the impedance Z of the circuit of Fig. 1 is

$$Z = \frac{1}{\mathrm{p}+\epsilon} + \nu \approx \frac{1}{\mathrm{p}+\epsilon} \tag{8}$$

while the admittance Y or $1/Z$ is

$$Y = \frac{1}{Z} = \frac{p + \epsilon}{1 + \nu p + \nu\epsilon} \approx \frac{p}{1 + \nu p} \qquad (9)$$

In equations (8) and (9), p denotes the conceptual differentiation operator d/dt, and $1/p$ means the conceptual integration operator with lower limit minus infinity, so that for all operands $h(t)$ that can be associated with a genuine physical situation (representable by means of a passive electrical network)

$$pp^{-1} h(t) = p^{-1}p\ h(t) = h(t) \qquad (10)$$

By operational calculus or otherwise, we can deduce from (8) that the formula for $V(t)$ when $I(t)$ is given is

$$V(t) = e^{-\epsilon t} \int_{-\infty}^{t} e^{\epsilon\tau} I(\tau)\ d\tau \qquad (11)$$

while that for $I(t)$, when $V(t)$ is given, is correspondingly

$$I(t) = \frac{1}{\nu}V(t) - \frac{1}{\nu^2}e^{-t/\nu}\int_{-\infty}^{t} e^{\tau/\nu} V(\tau)\ d\tau \qquad (12)$$

If $I(t)$ is the "admissible" function $h(t)$ first discussed in Section 2, $V(t)$ given by (11) has the Fourier transform

$$F(p)/(p + \epsilon) \text{ instead of } F(p)/p \qquad (13)$$

and thus for all finite p the change in this Fourier transform is arbitrarily small, but there is now a limit a/ϵ for the transform when p tends to zero, a being $F(0)$ which must be finite because $h(t)$ itself is admissible. Thus the process of "metrical integration" described by (11) gives an admissible output $V(t)$ for an admissible input $I(t)$. Again, if $V(t)$ is the same "admissible" function $h(t)$ the corresponding output $I(t)$ obtained from (12) has the Fourier transform $pF(p)/(1 + p\nu)$ instead of $pF(p)$. Hence when p tends to infinity, this Fourier transform tends to zero instead of to K and thus $I(t)$ is "admissible" as well as $V(t)$. Hence the process of "metrical differentiation" described by (12) also gives an admissible output $I(t)$ for an admissible input $V(t)$. Repeated "metrical" differentiation and/or integration will not destroy "admissibility" because the argument above for a single "metrical" differentiation or integration can be repeated indefinitely.

7. THE MATHEMATICIAN'S POINT OF VIEW

Given as before that $F(p)$ is the Fourier transform of $h(t)$ so that equations (5) apply, we note first that the infinite limits of integration indicate that a limit has been closed. The limiting values of $F(p)$ and $h(t)$ in (6a) and (6b) are necessary conditions for the integrals to converge, and imply that $h(t)$ has a smooth fine structure, as specified in Equation (7). Conditions (6a) and (6b) happen also to be the requirements that $h(t)$ shall be a function admissible for the representation of natural phenomena, although this fact is unimportant from the pure mathematician's point of view.

Now the limiting processes involved in equations (5) happen to be abnormally sensitive, and frequently abnormally difficult to evaluate in specific cases. There are very few analytic functions (of time) which are as well-behaved as $\exp(-k^2t^2)$. This noble expression has a nonzero value for any real finite value of t positive or negative (k being real) and tends to zero as t tends to $\pm\infty$; it has a Fourier transform. But most of the straightforward functions we encounter in our student days (hereinafter called "teaching functions" for brevity) fail to meet the requirement that they must be adequately specified for all values of time, not only those in which we happen to be interested. Thus the expression e^{at} (real part of $a > 0$) behaves excellently at its beginning when t is negative, but it has no finite end and is thus inadmissible unless "tailored" as explained below. Likewise the expression e^{-at} behaves excellently at its end when t is arbitrarily large and positive, but it has no finite beginning. An "admissible" function can be made up from "teaching functions" like this in much the same way as a tailor makes a garment from a roll of cloth. The "scissors" are the Heaviside step function $H(t - t_0)$ defined by

$$H(t - t_0) = 1(t > t_0) \qquad (14)$$

$$H(t - t_0) = 0(t < t_0) \qquad (15)$$

The value of $H(0)$ is not defined because it is not relevant. If we are ever under the delusion that we need to know it, the truth is that we have wrongly specified or formulated the problem under discussion. Thus the expressions

$$e^{at}H(-t) \text{ or } e^{-at}H(t) \qquad (16)$$

are admissible in the sense that they have both a beginning and an ending; their only defect is the abrupt discontinuity when $t = 0$. This discontinuity is merely a convenient artefact from the engineer's point of view. It is tacitly understood that any actual circuit contains at least residual elements which round off the theoretical discontinuity, so that a theoretically discontinuous input gives an output which may change extremely rapidly (e.g. when a switch is closed) but it is practice continuous and repeatedly differentiable. But from the mathematician's point of view, the "tailored" functions (16) cannot be differentiated. Other such functions cannot be integrated. Very few will

tolerate repeated differentiation and/or integration, for if $h(t)$ has a Fourier transform $F(\mathrm{p})$ which tends to zero as p tends to infinity and to a finite limit when p tends to zero, there is no guarantee that $\mathrm{p}^n F(\mathrm{p})$ or $\mathrm{p}^{-n} F(\mathrm{p})$ will do so.

This is a very severe limitation on the use of Fourier integrals, and has led mathematicians into various attempts to widen the applicability of the Fourier integral by means of Stieltjes or Lebesgue integrals and other less well-known techniques. It is here indicated, however, that the difficulty is due to what is really a schoolboy error of closing limits in the wrong order.

When a limit has to be taken which involves a "nominal" expression already in closed-limit form, it may be necessary to revert to the unclosed or "metrical" form of this expression to correct the order in which the limits are closed, and recover the information lost by the premature closure, as was in effect done when considering equations (3) and (4). Again, if

$$h(t) = \frac{1}{2\pi \mathrm{j}} \int_{-\mathrm{j}\infty}^{\mathrm{j}\infty} e^{pt} \frac{\mathrm{dp}}{\mathrm{p}} \qquad (17)$$

we have an indeterminate form, although it is given in many textbooks as $H(t)$. But

$$h_1(t) = \frac{1}{2\pi \mathrm{j}} \int_{-\mathrm{j}\infty}^{\mathrm{j}\infty} e^{pt} \frac{\mathrm{dp}}{\mathrm{p}+\epsilon} = e^{-\epsilon t} H(t) \qquad (18a)$$

$$h_2(t) = \frac{1}{2\pi \mathrm{j}} \int_{-\mathrm{j}\infty}^{\mathrm{j}\infty} e^{pt} \frac{\mathrm{dp}}{\mathrm{p}-\epsilon} = -e^{\epsilon t} H(-t) \qquad (18b)$$

and if the limits in (18*a*) and (18*b*) are closed by making ϵ tend to zero, the results differ by unity. If we try to evaluate the integral (17) by indenting the contour, leaving the imaginary axis for a small semicircle between the points $-\mathrm{j}\eta$ and $+\mathrm{j}\eta$, we obtain a result half way between, namely

$$h(t) = \tfrac{1}{2} \operatorname{sgn}(t) = \tfrac{1}{2}[H(t) - H(-t)] \qquad (19)$$

But by this process of indentation, what we have evaluated is not $h(t)$ given by (17) but

$$h_3(t) = \int_{-\mathrm{j}\infty}^{-\mathrm{j}\eta} e^{pt} \frac{\mathrm{dp}}{\mathrm{p}} + \int_{\mathrm{j}\eta}^{\mathrm{j}\infty} e^{pt} \frac{\mathrm{dp}}{\mathrm{p}} \qquad (20)$$

omitting the integral

$$h_4(t) = \frac{1}{2\pi \mathrm{j}} \int_{-\mathrm{j}\eta}^{\mathrm{j}\eta} e^{pt} \frac{\mathrm{dp}}{\mathrm{p}} = \int_{-\mathrm{j}\eta}^{\mathrm{j}\eta} \frac{\mathrm{dp}}{\mathrm{p}} \qquad (21)$$

which (for η sufficiently small to permit the factor e^{pt} to be replaced by unity) is $\pm\tfrac{1}{2}$. Alternatively, (17) is often expressed in terms of Dirichlet's integral

$$h(t) = \frac{1}{\pi} \int_0^{\infty} \frac{\sin \omega t}{\omega} \, d\omega = \tfrac{1}{2} \operatorname{sgn}(t) \qquad (22)$$

but here also we have omitted the critically important and indeterminate integral (21). Similarly the integral

$$h_5(t) = \frac{1}{2\pi \mathrm{j}} \int_{-\mathrm{j}\infty}^{\mathrm{j}\infty} e^{pt} \, \mathrm{dp} \qquad (23)$$

is indeterminate. On the other hand, if ν is positive, however small

$$h_6(t) = \frac{1}{2\pi \mathrm{j}} \int_{-\mathrm{j}\infty}^{\mathrm{j}\infty} e^{pt} \frac{\mathrm{dp}}{1+\nu \mathrm{p}} = \frac{1}{\nu} e^{-t/\nu} H(t) \qquad (24)$$

$$h_7(t) = \frac{1}{2\pi \mathrm{j}} \int_{-\mathrm{j}\infty}^{\mathrm{j}\infty} e^{pt} \frac{\mathrm{dp}}{1-\nu \mathrm{p}} = -\frac{1}{\nu} e^{t/\nu} H(-t) \qquad (25)$$

without any difficulty or ambiguity at all.

8. CONCLUSIONS

From the point of view both of the engineer and of the pure mathematician, difficulties encountered in calculating with ideal, nominal or quantized quantities expressed in closed-limit forms are easily evaded with no loss of precision by replacing conceptual entities in closed-limit forms by the corresponding metrical or open-limit forms. If this is done, there is no need to study techniques involving anything more than straightforward differentiation and ordinary Riemann integration, and no discontinuity more complicated than an isolated step need be considered.

9. ACKNOWLEDGEMENT

The author is greatly indebted to Mr. C.G. Mayo, without whose collaboration, criticism, encouragement and help this report could never have been written.

Printed by BBC Research Department, Kingswood Warren, Tadworth, Surrey